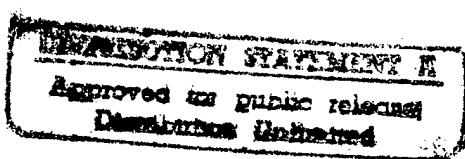


AFOSR-TR-97-0155

Individual Differences in Skilled Performance Errors

AASERT Grant F49620-93-0378
Air Force Office of Scientific Research
Life Sciences Directorate
Bolling AFB DC 20332-6448

Final Report
1 July 93 - 30 Nov 96



Dan J. Woltz
Michael K. Gardner DTIC QUALITY INSPECTED 4
Brian G. Bell

University of Utah

19970407 013

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

Public reporting burden for this collection of information is estimated to be no 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188), Washington, DC 20503.

1. AGENCY USE ONLY (Leave blank)	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED	
	28 Feb 97	Final Report 1 July 93 - 30 Nov 96	
4. TITLE AND SUBTITLE		5. FUNDING NUMBERS	
Individual Differences in Skilled Performance Errors		F49620-93-0378	
6. AUTHOR(S)		61103D	
Dan J. Woltz, Michael K. Gardner, & Brian G. Bell		3484/YS	
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)		8. PERFORMING ORGANIZATION REPORT NUMBER	
University of Utah Salt Lake City, UT 84112		Final Report	
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)		10. SPONSORING/MONITORING AGENCY REPORT NUMBER	
AFOSR/NL Bolling AFB, DC 20332-6448			
11. SUPPLEMENTARY NOTES			
12a. DISTRIBUTION/AVAILABILITY STATEMENT		12b. DISTRIBUTION CODE	
Unlimited			
13. ABSTRACT (Maximum 200 words)			
<p>Three studies were conducted to investigate questions concerning which type of individual is most susceptible to errors in skilled (highly practiced) cognitive task performance. The first study was a small pilot project conducted with university students. The remaining two studies were larger scale projects conducted with Air Force enlisted personnel. The research findings suggested that self-reported error-proneness is not related to performance in at least one class of cognitive skill tasks or to working memory task performance. Instead, measures of error-proneness appear to reflect how a person perceives themselves or how they wish to present themselves. The research findings also suggested that measures of working memory capacity (i.e., measures of general mental resource availability or mental workspace) were predictive of undetected errors in skilled performance after extensive skill practice. This finding suggests the need for further research as it has potential implications for selection of individuals to critical jobs. This finding also has implications for current theories of individual differences in learning. Finally, the research produced inconsistent findings regarding the relationship of anxiety to cognitive performance errors. More research is needed to resolve questions in this latter topic area.</p>			
14. SUBJECT TERMS		15. NUMBER OF PAGES	
Individual differences, cognition, human error, skilled performance, human memory		39	
16. PRICE CODE			
17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
unclassified	unclassified	unclassified	UL
NSN 7540-01-280-5500			

Contents

Introduction	3
Study 1: Personality Variables and Skilled Performance Errors.....	3
Method	4
Measures.....	4
Apparatus	5
Subjects	5
Procedure.....	5
Results and Discussion.....	5
Number Reduction Training Data.....	5
Number Reduction Transfer Data	5
Correlations with Self-report Measures	6
Study 2: Working Memory, Anxiety, and Self-reported Error-proneness.....	8
Method	9
Subjects	9
Apparatus	9
Measures.....	9
Procedure.....	11
Results	12
Number Reduction Mean Data.....	12
Working Memory Task Mean Data	15
Descriptive Data for Self-report and Speed Measures	15
Correlations	16
Multiple Regression Analyses.....	19
Discussion	20
Study 3: Anxiety and Working Memory.....	21
Method	22
Subjects	22
Apparatus	22
Measures.....	22
Procedure.....	23
Results	23
Number Reduction Mean Data.....	23
Working Memory Task Mean Data	26
Descriptive Data for Self-report Measures.....	27
Correlations	27
Multiple Regression Analyses.....	31
Discussion	32
Conclusions & Implications	33
References	35
Students Supported	37
List of Publications.....	38
List of Professional Presentations	39

Introduction

Over the last century, a considerable body of research has examined the benefits of practice on skilled performance (e.g., Anderson, 1983, 1987; Bryan and Harter, 1899; Crossman, 1959). In contrast, few studies have looked at the errors made by skilled individuals, particularly with the idea that highly practiced individuals may be more prone to certain types of performance errors than less practiced individuals. Although skilled performers rarely make mistakes in most conditions, human error by highly skilled individuals does occur and it can have disastrous consequences in many settings.

Research supported by the parent grant (F49620-93-1-0094) found that skilled performers are more susceptible than less skilled performers to some error types during the performance of sequential skills (i.e., cognitive skills that require the proper sequencing of a set of component operations). Specifically, skilled performers tend to make undetected mental errors when they are required to use novel sequences of operations that have not been previously practiced (i.e., highly practiced operations in a new sequence). These errors have been termed *strong-but-wrong* errors (Reason, 1990). It appears as though strong memory representations for processing sequences that have been practiced extensively fire inappropriately when new sequences resemble them in the initial components (Woltz, Bell, Kyllonen, & Gardner, 1996; Woltz, Gardner, & Bell, 1995). Of importance, subjects tended to be unaware of these errors, thus precluding self-correcting actions. Errors in the real world of this kind are necessarily dangerous when they go undetected by the performer.

The research supported by this AASERT grant examined individual differences in this type of performance error made after subjects acquire a moderately high level of skill on a cognitive task. After determining that individual differences in skilled performance errors were reliable (i.e., stable over occasions and conditions), we investigated whether various ability and personality characteristics of individuals predicted error-proneness. In addition, the research attempted to determine the mechanisms by which one of the individual difference predictor variables, namely anxiety, may operate.

Three research studies were conducted on individual differences over the three year period. Study 1 was a small pilot study that investigated the relationship between a set of personality variables, including self-reports of error-proneness, and skilled performance errors. Study 2 further investigated self-report of error-proneness as well as self-report of anxiety and performance measures of working memory capacity and processing speed. Studies 1 and 2 were also reported in Bell, Gardner, & Woltz (in press). Study 3 primarily investigated measures of anxiety and working memory capacity to test theories of how anxiety might affect skilled performance. The details of these studies will be reported in sequence.

Study 1: Personality Variables and Skilled Performance Errors

Study 1 focused on the relationship between self-report inventories of error propensity and actual error making on a cognitive skill task. The skill task used was *number reduction*, which was a modification of a task originally developed by Thurstone and Thurstone (1941). In this task, subjects are taught a set of four rules for reducing four digit number strings to a single digit response. Subjects apply the rules to pairs of digits, proceeding from left to right. The application of each rule yields a single digit answer to the pair of digits. This digit becomes the first digit in the next pair. Processing proceeds until only a single digit remains.

The task is best understood by example. Consider the string: 4568. The four rules subjects learn are: (a) the *same* rule, which states that if two digits are the same (e.g., 55), the answer is that same digit (5); (b) the *contiguous* rule, which states that if two digits begin an ascending or descending series (e.g., 67 or 43), the answer is the next digit in that series (8 or 2); (c) the *midpoint* rule, which states that if two digits differ by two (e.g., 35), the answer is the digit midway between them (4); and (d) the *last* rule, which states that if two digits differ by more than two (e.g., 38), the answer is the latter of the two digits (8). Our example (which represents the CONTIGUOUS-SAME-MIDPOINT rule sequence) would be solved as follows: (45) = 6. Six becomes the first digit in the next pairing, (66) = 6. The final pairing is therefore (68) = 7, and the subject would respond by pressing "7". Note that the subject does not input the intermediate responses; instead these operations are performed in the subject's head.

Each four digit stimulus string therefore requires the application of three rules for its solution. Strings are designed so that no rule appears more than once per string, and that over strings (within an individual) the frequency of occurrence of all rules in all serial positions is balanced.

To produce *strong-but-wrong* errors, we trained subjects on 8 of the 24 possible sequences of three (non-duplicating) rule applications (e.g., Woltz, Bell, Kyllonen, & Gardner, 1996; Woltz, Gardner, & Bell, 1995). Subjects saw many (i.e., 15) exemplars of each rule sequence (e.g., 4568, 9875, and 2346 are three exemplars of the CONTIGUOUS-SAME-MIDPOINT rule sequence). This training lasted for several sessions, and subjects developed considerable skill at the task (i.e., solution latencies displayed a large reduction that follows the power law of learning [Newell & Rosenbloom, 1981]). During a transfer session, we presented subjects with problems using the entire set of 24 rule sequences. Subjects were generally not aware that any change had occurred, as the surface structure of items appeared quite similar to training; however, it was these new rule sequences that produced the vast majority of undetected errors.

During the transfer session we also asked subjects to achieve 100% accuracy while going as fast as possible. We informed them that they could retake any trial they thought they had made an error on by pressing the space bar on the computer. They were told that this was done to help them reach their goal of 100% accuracy, but, in fact, it allowed us to separate those trials on which subjects made a detected error from those trials on which subjects made an undetected error.

Method

Measures. Two self-report measures of error making were included in Study 1. The first was the *Cognitive Failures Questionnaire* (CFQ; Broadbent, Cooper, Fitzgerald, and Parkes, 1982), which assesses slips and lapses in memory, perception, and action. Due to an error, only 22 of the 25 items on the CFQ were used in the current study. Since 88% of the items were used, we doubt that the three excluded items would have seriously changed our results. The second self-report measure was the *Error-Proneness Questionnaire* (EPQ; Reason & Mycielska, 1982). It asks subjects how often they make various kinds of slips and lapses in everyday life.

We included the Desirability scale from the *Jackson Personality Research Form* (Jackson, 1989). We included this scale because it seemed plausible that subjects who wished to be seen in a socially desirable light might modify their responses (either consciously or unconsciously) to questionnaires inquiring about their error-proneness. Also given were the Dickman Impulsivity Scale (Dickman, 1990), which is a measure of impulsivity, and several additional scales from the Jackson Personality Inventory (Jackson, 1976) and the Personality

Research Form (Jackson, 1989), which measure a variety of constructs, including Endurance, Impulsivity, and Achievement Motivation.

Skill task and apparatus. The cognitive skill task was number reduction, which was described above. Subjects performed the skill task on IBM compatible microcomputers with standard keyboards and SVGA monitors. Materials were presented in 24 X 80 text mode. The software was written to achieve millisecond timing of response latency, and to record detected and undetected errors (Walker, 1985).

Subjects. The subjects were 43 University of Utah students, of which 24 were female and 19 were male. All subjects completed both self report error measures and the other personality scales; however, only 29 of the 43 subjects completed the cognitive skill task.

Procedure. Subjects participated in five experimental sessions over a two week period. Sessions were scheduled such that no two sessions were conducted on the same day, and no more than four days intervened between sessions. Each session lasted less than one hour.

During the first four sessions, subjects received training on 8 of the 24 possible three-rule sequences. These eight were randomly chosen for each subject, with the constraint that each rule had to be used with equal frequency in each serial position. During session one, subjects solved 10 blocks of 24 trials each; during sessions two, three, and four, subjects solved 15 blocks of 24 trials each. At the beginning of each training session, subjects were provided with examples of each of the component rules and received four practice trials with each rule (16 total). They also saw examples of rule sequences and received three practice sequences. At the end of sessions one, three, and four subjects also filled out the paper and pencil measures.

During training on number reduction, feedback was provided to discourage errors and increase response speed. Latency feedback was presented for 1 sec following correct responses, and the word WRONG with a low tone was provided for 2 sec following incorrect responses. After each block of trials, the percent correct and median latency for that block was presented along with instructions for performance on the next block. If the subject's error rate was 15% or higher, they were told to slow down to reduce their errors. With an error rate of 5% or less, they were instructed to respond more quickly. Every subject was told to go faster than they had in the previous block and get approximately 90% of the trials correct. Following this instruction, they were shown the median latency for each previous block.

The fifth session comprised the transfer portion of the experiment. Subjects solved 18 blocks of 24 trials each. All 24 rule sequences were used during transfer: 8 old and 16 new. In contrast to training in which subjects were told to achieve 90% accuracy, during transfer subjects were told to achieve 100% accuracy while going as fast as possible. Subjects were instructed to press the spacebar to retake any trial on which they thought they had made an error. They were instructed that corrected error trials would not count against their performance criteria of 100% accuracy. No accuracy feedback was provided during transfer. Subjects did, however, receive latency feedback at the end of each transfer block.

Results and Discussion

Number reduction training data. As Figure 1 shows, latency declined with practice, with a power function fitting the data ($R^2 = 0.98$). Subjects also achieved their performance goal of approximately 10% errors during training.

Number reduction transfer data. We used the nonparametric Wilcoxon matched pairs z -test to test for differences in error rates, due to the fact that the distribution of errors was nonnormal. Latency differences, however, were tested using the paired sample t -test because these distributions approached normality. During transfer, subjects had a mean detected error rate of

2.55% ($SD = 2.67\%$, Skewness = 1.55) and a mean undetected error rate of 8.60% ($SD = 7.51\%$, Skewness = .84; Wilcoxon $z = 3.36$ $p < .01$). As in previous research, there was a significantly higher undetected error rate on the new sequences ($M = 9.59\%$, $SD = 9.12\%$, Skewness = 1.09) in contrast to the old ($M = 6.63\%$, $SD = 5.63\%$, Skewness = .93; Wilcoxon $z = 2.40$ $p < .05$). The higher undetected error rate among the new sequences demonstrates that our transfer condition was effective in producing undetected errors. The mean latency during transfer for new sequences was 2098 ms ($SD = 416$ ms, Skewness = -.08), while the mean latency during transfer for old sequences was 1963 sec ($SD = 427$, Skewness = .06; $t(28) = -5.05$, $p < .001$). Thus new sequences, which produced more undetected errors, also produced significantly longer latencies.

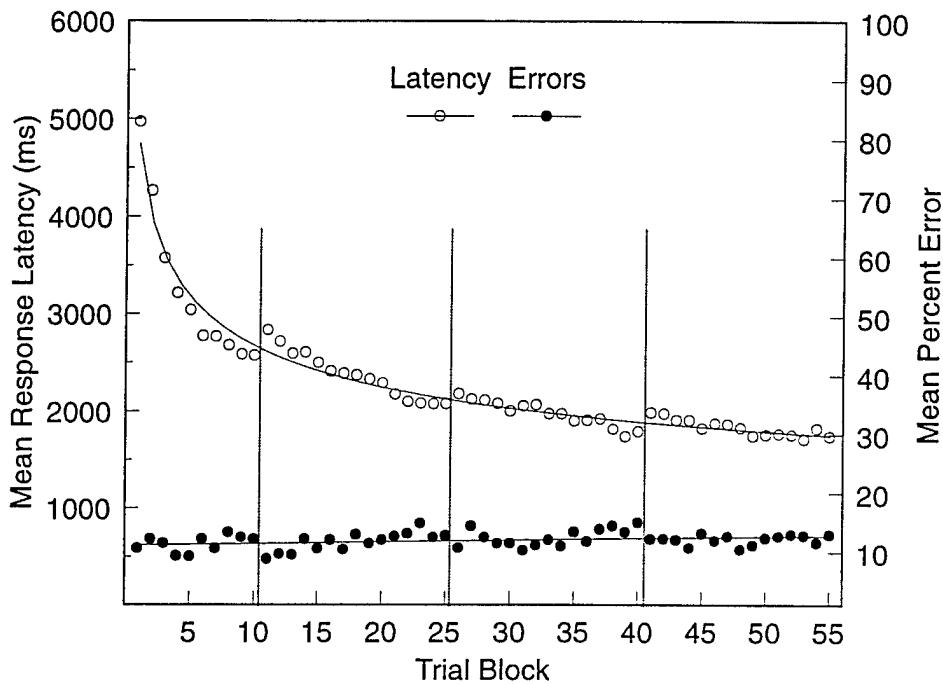


Figure 1. Mean latency and error data for training blocks of Study 1.

Correlations with self-report measures. Table 1 presents a subset of the correlational data from Experiment 1. The impulsivity, endurance, and achievement motivation measures were excluded. Although the omitted individual difference measures were significantly intercorrelated, none of them were correlated beyond chance levels with skilled performance errors.

Several things are apparent from the data reported in Table 1. First, detected and undetected errors were correlated near zero with each other, and this correlation was not statistically significant. This suggests that subjects were not simply varying their criteria for using the spacebar, thus switching errors from the detected category to the undetected category. Second, responses to the CFQ and the EPQ were significantly related to each other, and this correlation was the largest in Table 1. Thus, these self-report measures of error-proneness seem to be measuring something in common. Third, the CFQ and EPQ were not significantly correlated with either undetected or detected errors. It seems that self-report measures of error-proneness do not reflect actual error making in this performance situation. It should be pointed out that both the CFQ and EPQ are broad measures of error making; they inquire about errors in

numerous tasks and domains. Errors in the number reduction task obviously involve a very specific situation. Thus, the failure to find significant correlations does not completely invalidate the instruments. On the other hand, the number reduction task was designed to index a significant type of skill: the application of sequential cognitive operations while storing results mentally. The failure of these questionnaires to correlate significantly with errors in this environment is a noteworthy finding.

Table 1
Means, Standard Deviations, and Correlations Among Dependent Variables (from Transfer Session) in Study 1^a

	M	SD	1	2	3	4	5	6
1. Latency of Correct Responses (ms)	2052.90	413.88	1.00	-.15	-.52	.13	.18	-.02
2. Undetected Errors (%)	9.59	9.12		1.00	.06	-.16	-.03	.26
3. Detected Errors (%)	2.55	2.60			1.00	-.05	.17	.20
4. Cognitive Failures Questionnaire	2.60	0.42				1.00	.64	-.41
5. Error-Proneness Questionnaire	5.27	1.54					1.00	-.22
6. Jackson Social Desirability Scale	68.90	17.55						1.00

^a Correlations involving errors (either undetected or detected) based on n = 29; r > .31 is significant at p=.05 . Correlations not involving errors based on n = 43; r > .25 is significant at p=.05.

Finally, at least one of the self-report measures (CFQ) was significantly related to a measure of social desirability, consistent with our hypothesis that self report measures are influenced by social desirability. Whether the source of this relationship was truly "social desirability" is difficult to say. Measures of social desirability are not well understood, and the source of the relationship may not be social desirability per se, but rather some other factor partially assessed by the social desirability scale we chose to use.

One last finding of interest is the significant negative correlation between latency of correct responses and detected errors. The correlation between latency of correct responses and undetected errors is also negative, although it is not statistically significant. These correlations reflect the fact that those individuals who achieved higher levels of skill (i.e., they had small correct response latencies) made more errors, a result that is unusual in the skill acquisition area. It is, however, consistent with the hypothesis that individuals become susceptible to certain types of errors (e.g., strong-but-wrong errors) only after they become reasonably skilled performers (i.e., certain "strong" response tendencies can only exist after a sufficient degree of training).

Study 2: Working Memory, Anxiety, and Self-reported Error-proneness

We concluded from Study 1 that self-report measures of error-proneness may tell us more about individuals' perception of themselves than about their actual performance on a skilled cognitive task. Clearly, larger numbers of subjects and greater power would have made it easier to assess the actual magnitude of relationships among the measures in Study 1. It seemed necessary to replicate these finding with a larger sample. Study 2 provided this replication by including the CFQ.

Study 2 had two additional goals. First, we wanted to examine the relationship between processing resources, as indexed by working memory, and undetected errors in skilled performance. It seemed to us that the ability to recognize transfer task demands that require unfamiliar processing sequences would require a subject to devote a certain amount of working memory capacity to monitoring performance. Subjects who had less working memory capacity to begin with should be unable to free up sufficient resources for this task, leading to undetected errors. Thus Study 2 employed two measures of working memory capacity, which were performance based. We also included two measures of simple processing speed. Verbal and quantitative processing speed measures assessed processing efficiency of simple mental operations similar to those used in the number reduction task. As with working memory, it was hypothesized that more efficient processing (i.e., lower latencies) would be associated with fewer undetected errors.

Our second additional goal in Study 2 was to investigate self-reported anxiety as a potential predictor of skilled performance errors. Two theories have been proposed to explain why anxious subjects could perform more poorly on a variety of tasks. One theory focuses on the working memory limitations of anxious subjects and states that anxiety reduces working memory capacity, thereby impairing the performance of anxious subjects on complex tasks (Eysenck, 1983). In effect, anxious subjects are thought to be in a dual-task situation and their performance suffers. Especially with more difficult tasks which require more capacity, anxious subjects supposedly perform more poorly than nonanxious subjects because the combination of task difficulty and anxiety-related thoughts will exceed their working memory capacity. The second theory, which provides an attentional explanation, states that anxious subjects ignore relevant cues (Eysenck, 1983). According to this theory, anxious subjects perform more poorly on complex tasks because these tasks contain a larger number of relevant cues.

The working memory theory and the cue utilization theory tend to make the same general predictions about the performance of anxious subjects, namely, that anxious subjects will perform more poorly on complex or difficult tasks. However, we believe that these theories make different predictions about the performance of anxious subjects under certain conditions. The primary focus of our investigation of anxiety was to examine whether working memory limitations are responsible for the hypothesized poorer performance of anxious subjects, whether the cue utilization theory provides a better explanation, or whether both theories account for the behavior of anxious subjects in skilled performance.

Recall that in Study 1, subjects performing the number reduction task were exposed to partial-match sequences that started with the same first two rules as the old sequences, and mismatch sequences that started with an unfamiliar two-rule sequence. This study also used partial-match, mismatch, and old sequence trials, but in addition, the partial-match and old trials started with the same first three digits as well as the same first two rules (e.g., 3499 and 3497 differ only in the last rule and the last digit). The physical similarities between the partial-match items and the old items should have produced errors more errors in anxious subjects if they do

not attend to the relevant cue, which is the last digit. If anxious subjects perform more poorly on these items than on mismatch items, this would be consistent with the cue utilization theory of how anxiety affects performance.

In contrast, the working memory theory predicts that anxious subjects should perform more poorly on the items that produce the heaviest cognitive load, namely the mismatch items. A skilled memory representation may be used to respond to the partial-match items because these items are similar to the old items, but with the mismatch items, the subjects must return to more rudimentary representations of component rules. That is, the mismatch items contain unfamiliar sequences, so subjects must solve each item by using the component rules in a stepwise fashion. This presumably places greater demands on working memory than the partial-match trials.

This study tested a working memory explanation in three ways. First, subjects performed two separate working memory tasks: one that presented verbal stimuli and one that presented numbers. These tasks required the subject to remember either a list of numbers or words and transform the stimuli. Anxious subjects should perform more poorly on these tasks if they have less working memory capacity available due to worry or other interference. The role of working memory limitations was also tested by adding a working memory load to the skill acquisition task. A concurrent load was added to every other block of transfer trials. Subjects had to keep track of how many times they pressed certain number keys as they responded to the items. If anxious subjects have less capacity, then they should be more adversely affected by the introduction of this load than non-anxious subjects. Finally, working memory limitations were examined by introducing new sequence trials during the transfer phase of the experiment. These were the mismatch sequences that were described above. If anxious subjects have less capacity then they should perform more poorly on these items than on the partial-match items.

It should be pointed out that anxiety may not always lead to poorer performance. As Eysenck (1984) noted, anxious subjects may compensate by expending more effort. However, this issue is not addressed by the theories that are tested in this research, so this possibility will not be discussed further.

Method

Subjects. The subjects were 78 US Air Force enlisted personnel who had almost completed basic training at Lackland Air Force Base, TX. Four of these subjects were eliminated from the study because their performance indicated a lack of effort. These subjects had a high error rate on practice items after extensive training. Another five subjects were removed from the analysis because data was missing for some of the transfer blocks of the skill acquisition task. All of the subjects in the study were male. The age of Air Force recruits ranges from 17 to 27.

Apparatus. The experimental task was administered on Zenith Z-248 microcomputers with standard keyboards and EGA color video monitors. Materials were presented on the monitors in 24 X 80 text mode. Software was written to achieve millisecond timing of response latency recording (Walker, 1985).

Measures. The subjects received computerized versions of two anxiety measures: the *Test Anxiety Inventory* (TAI; Spielberger, Gonzalez, Taylor, Anton, Algaze, Ross, and Westberry, 1980) and the *State-Trait Anxiety Inventory* (STAII; Spielberger, Gorsuch, Lushene, Vagg, and Jacobs, 1983). A computerized version of the *Cognitive Failures Questionnaire* (CFQ; Broadbent, Cooper, Fitzgerald, and Parkes, 1982) was used to measure the tendency to make cognitive errors.

Two tasks from the Air Force CAM 4 computerized test battery (Kyllonen, Woltz, Christal, Tirre, Shute, & Chaiken, 1990) were used to measure working memory capacity. The measures,

one of which was quantitative and the other verbal, consisted of 32 items. Each item presented subjects with four to seven stimuli, and required them to remember the last three stimuli in order. The verbal stimuli were one-syllable adjectives such as big, cold, and fast. Numerical stimuli consisted of the digits 1 through 9. Verbal stimuli were presented at the rate of one word every 2.5 seconds, and numerical stimuli were presented at the rate of one number every 1.25 seconds. The color of the stimuli was varied so that subjects had to transform the words or numbers. When the stimulus was white, the subject did not have to transform the word or number. However, when the stimulus was yellow, the subject had to subtract the number that was presented from 10 (e.g., $10 - 6 = 4$) or remember the antonym of the word (e.g., the opposite of big is small). The last three stimuli for an item were either consistent, either all white or all yellow, or were inconsistent, which means that white and yellow stimuli were mixed. Subjects had 18 seconds to respond to each item. These tasks presumably tapped working memory capacity because the subjects were required to retain and manipulate the stimuli.

The subjects responded after all stimuli had been presented by selecting from eight alternatives. Subjects also used a three-point scale to provide confidence ratings concerning the correctness of their responses. The three points on the scale were labeled "I think I got it wrong", "I am not sure", and "I think I got it right". After each block of eight trials, the subjects were provided with accuracy feedback. The feedback also indicated how many items the subject thought they had answered correctly.

Two tasks that are not published, one verbal and one quantitative, measured how quickly the subject could perform simple mathematical calculations (e.g., $6 - 5$) or decide whether simple words had similar meanings or different meanings (e.g., argue and debate). Subjects responded by typing "L" on the computer if the words were synonyms or the second number was the answer, and "D" if not. Subjects received latency feedback at the end of each trial and accuracy feedback at the end of each block. The measure of performance was response speed.

As noted in Study 1, the number reduction task requires that subjects use a total of four rules to reduce a set of numbers to a single digit. On any given trial, the subject uses only three of the four rules. Each rule sequence is represented by many instances, with each instance using a different set of numbers, but the same set of rules. For example, the instance 1 2 3 6 represents the sequence CSL (Contiguous Same Last are the rules for this sequence).

The practice phase consisted of 20 blocks with 32 trials per block. The practice blocks contained a total of 16 unique instances, with four instances representing each of four rule sequences. The transfer phase consisted of two warm-up blocks, that contained 32 practice items each, and eight transfer blocks. Each of the eight transfer blocks contained 16 practice, 8 mismatch, and 8 partial-match trials. The transfer blocks contained the 16 original practice instances as well as 16 new instances that represented 4 partial-match sequences and 16 new instances that represented 4 mismatch sequences. The practice and partial-match sequences were counterbalanced so that the 4 practice sequences for half of the subjects were used as the 4 partial-match sequences for the other group. All of the sequence types (practice, partial-match, and mismatch) used each of the four component rules (Contiguous, Last, Midpoint, and Same) an equal number of times in each position.

Half of the transfer blocks contained a working memory load. Subjects were required to keep track of how many times they pressed a pair of keys (such as 1 and 4) while they performed the task. They were provided with feedback at the end of the block which told them how many times they had actually responded with the target keys. For each working memory block, a different pair of response keys was chosen to be remembered.

The partial-match trials started with the same two rules and the same three digits as practice trials. For example, if the subject received the sequence CSL, which is represented by the instance 1 2 3 6, in practice, then they would see the sequence CSM in transfer, and the instance 1 2 3 5 would represent this sequence. Notice that the practice and partial-match instances differ on the last digit, and these digits are adjacent (5 and 6). By using adjacent digits, it was hoped that the stimuli would be more confusable. The mismatch items started with a different 2-rule sequence. For example, if CSL was a practice sequence and CSM was a partial-match sequence, then either CLM or CML would be the mismatch sequence. Mismatch instances and practice instances that started with the same rule differed on the first digit as well. If 1 2 3 6, which represents the sequence CSL, was a practice instance, then a mismatch instance that started with the Contiguous rule (C) could not start with a 1. Mismatching on starting sequence and first digit should increase the working memory load. Items for a given mismatch sequence all started with a different digit as well to increase the load on working memory. All of the practice instances for a given sequence started with a unique digit so that only the practice instance and corresponding matching instance shared the same first three digits. By doing this, it was hoped that subjects during training would not become familiar with items that started with the same three digits but ended in a different digit.

During transfer, subjects could retake the previous item by pressing the spacebar if they thought they had made an error. This was done to separate detected and undetected errors.

Procedure. The subjects were tested in groups of 30-40, with each subject at an individual testing carrel containing a microcomputer. Each subject participated for 3.5 hours in a series of experimental tasks.

At the beginning of each session, the subjects were given a general orientation to the experiment and a few minutes of practice locating keys on the computer keyboard. All of the instructions were computer administered, and proctors were available to answer questions. Following the orientation, the subjects completed a series of tasks, which are outlined below.

At the beginning of the experiment, subjects were provided with examples of each of the four component rules that are used in the number reduction task and received four practice trials with each rule. They also saw examples of rule sequences and received three practice sequences. The subjects then took 10 practice blocks of number reduction, with each block consisting of 32 items. Either the STAI or the TAI, one of the working memory tasks, and one of the speed measures were completed next. Another 10 practice blocks of number reduction followed, with a different anxiety measure, working memory task, and speed measure being presented. Subjects finished the session by taking the transfer blocks of number reduction, which consisted of 10 blocks of 32 items, and the Cognitive Failures Questionnaire. Eighteen subjects also had their heart rates recorded while they took the number reduction task, the State Trait Anxiety Inventory, and the working memory measures.

During the practice blocks of number reduction, feedback was provided to discourage errors and increase response speed. Latency feedback was presented for 1 sec following correct responses, and the word WRONG and a low tone was provided for 2 sec following incorrect responses. After each block of trials, the percent correct and median latency for the block was presented along with instructions for performance on the next block. If the subject's error rate was 15% or higher, they were told to slow down to reduce their errors. With an error rate of 5% or less, they were instructed to respond more quickly. Each subject was told to go faster than in the previous block and get approximately 90% of the trials correct. Following this instruction, they were shown the median latency for each previous block.

During the transfer phase of the number reduction task, subjects were told to achieve 100 % accuracy while going as fast as possible. Subjects were instructed to press the spacebar to retake any trial on which they thought they had made an error. They were instructed that corrected error trials would not count against their goal of 100% accuracy. No accuracy feedback was provided during transfer, but subjects did receive latency feedback at the end of each transfer block.

Results

Number Reduction Mean Data. The learning curves (Figure 2) reveal that subjects generally conformed with the performance goal of 10% errors during training. Furthermore, latency declined with practice, and a power function provided a close fit to the data. Finally, when the subjects were asked to detect and correct their errors (Blocks 21 and 22), their error rate declined while latency increased.

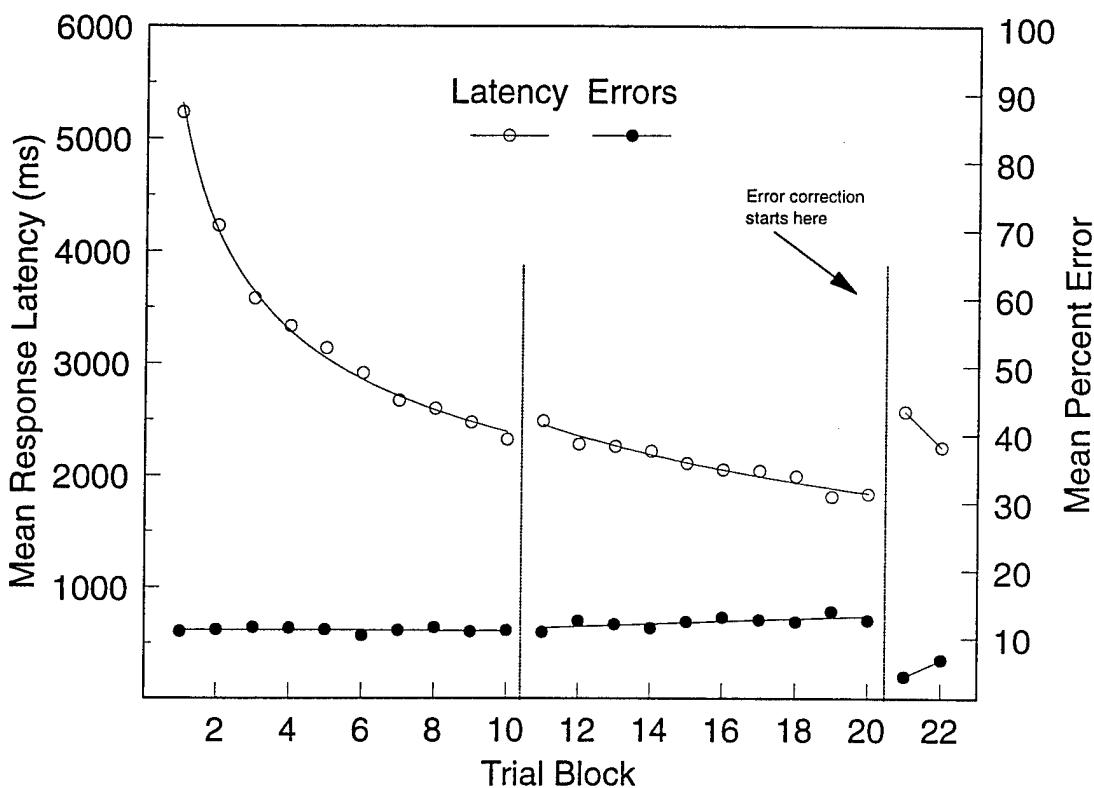


Figure 2. Mean latency and error data from training blocks of Study 2.

Figure 3 shows the mean latency of correct responses in transfer trials by trial type, block, and load condition. It reveals that subjects responded more quickly with old (practice) items but that the latency of correct responses for partial-match and mismatch items did not differ. The results from an analysis of variance confirmed these observations. A significant difference in the latency of correct responses was found between the old and new items, $F(1,68) = 136.91$, $MSe = 4688729$, but not between the partial-match and mismatch items, $F(1,68) = 1.33$, $p > .05$. If the mismatch trials taxed working memory capacity, then the latency of correct responses for these

items should have been greater. Since the latency of correct responses was greater for both types of new items compared to the old items, it appears that both types consumed working memory capacity to the same degree. The significant interaction between load and item type, $F(1,68) = 6.12$, $MSe = 1491471$, shows that subjects were slower in making correct responses for new items when a load was introduced. This suggests that the combination of a load and unfamiliar items may have taxed the subjects' working memory capacity.

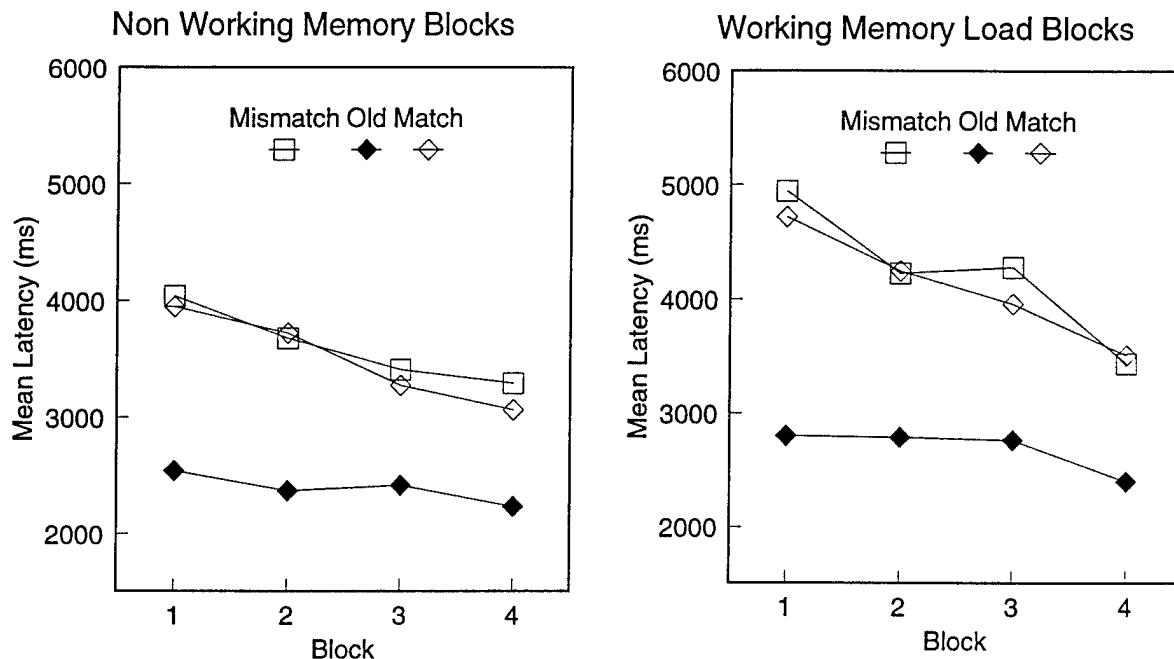


Figure 3. Mean latency for transfer blocks by working memory condition in Study 2.

Figure 4 shows the mean percentage of errors in transfer trials by trial type, block, and load condition. Only undetected errors were analyzed because few errors were detected, and because earlier work with this task primarily examined undetected errors. The mean percentage of undetected errors was 23.4% ($SD = 21.1\%$) while the mean percentage of detected errors was 3.08% ($SD = 3.68\%$, $F(1,68) = 58.33$, $MSe = 244.07$). Figure 4 reveals that subjects made fewer errors with old items, $F(1,68) = 57.03$, $MSe = 1110$, and there was a difference in accuracy between the types of new items. The percentage of errors was greater for the partial-match items compared to the mismatch items, $F(1, 68) = 11.05$, $MSe = 382$. The working memory load did not increase the number of errors, $F(1,68) < 1$.

We analyzed the response latency of errors to check the assumption that partial-match trials would produce fast slips (strong-but-wrong errors). As shown in Figure 5, the latency of errors for the partial-match ($M = 3114$, $SD = 2018$) and mismatch items ($M = 3300$, $SD = 2270$) did not differ ($F(1,63) = 2.00$ $MSe = 1759348$). This finding does not support the notion that errors on the partial-match items were due to a failure to recognize the difference between partial-match items and old items on the last digit.

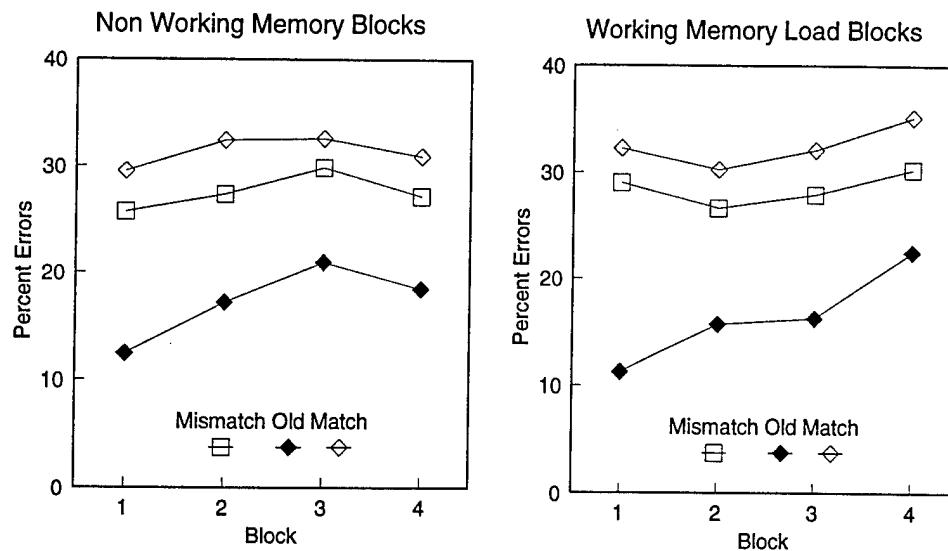


Figure 4. Mean Error rate for transfer blocks by working memory condition in Study 2.

In contrast to the findings from Woltz et. al.(1995), in which mismatching items appeared to consume more working memory capacity than partial-match items, and the latency of errors for mismatching items and partial-match items differed, the two new item types did not have the desired effect on performance in this study. Although mismatching items may have consumed more working memory capacity than old items, as measured by the latency of correct responses, there was no difference between the latency of correct responses for partial-match and mismatch items. Furthermore, the latency of errors for partial-match items did not differ from the latency of errors for mismatch items. This suggests that errors on partial-match items may not be due to the failure to attend to the last digit.

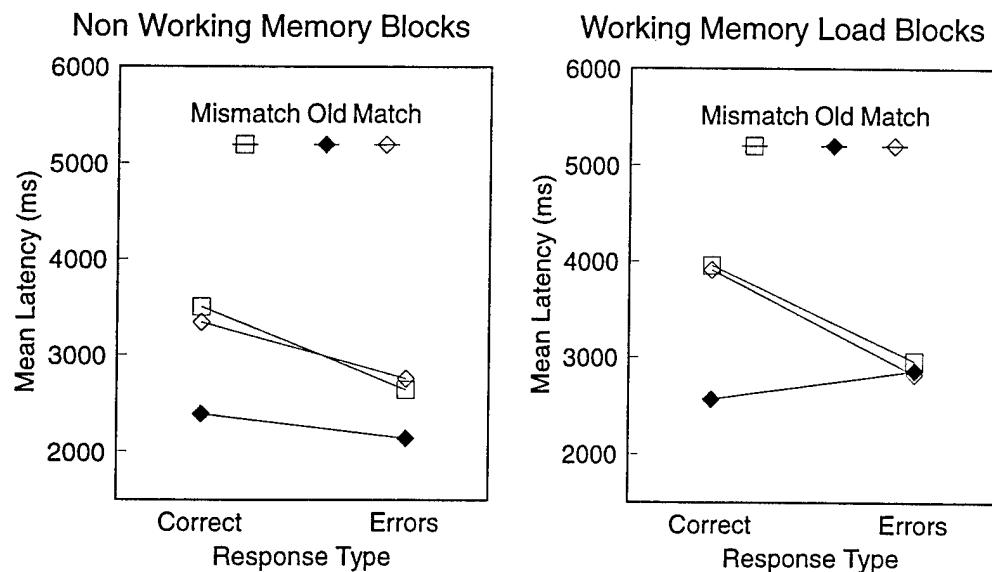


Figure 5. Mean latency for correct trials and undetected errors in Study 2.

Working Memory Task Mean Data. The effects of listlength, color of the stimuli (i.e., the need to transform the stimuli), and consistency of color were examined. These working memory manipulations appeared to have the desired effect. Accuracy declined as listlength increased for both the verbal and quantitative working memory tasks (Figure 6). The effects of listlength were significant for the verbal, $F(3,66) = 24.61$, and quantitative, $F(3,66) = 21.48$, tasks. There was a significant interaction of color and consistency, ($F(1,68) = 8.10$, $MSe = 940.36$) for the verbal working memory task, and the effects of both color, $F(1,68) = 44.42$, $MSe = 783.91$) and consistency, $F(1,68) = 5.67$, $MSe = 517.5$) were significant for the quantitative working memory task. Overall, subjects performed better with white stimuli, which did not require any transformations. Also, subjects generally performed better when all of the stimuli were the same color. The mean percentage of errors on the verbal working memory task was 38.32% ($SD = 22.67$) and on the quantitative working memory task was 19.66% ($SD = 18.15$).

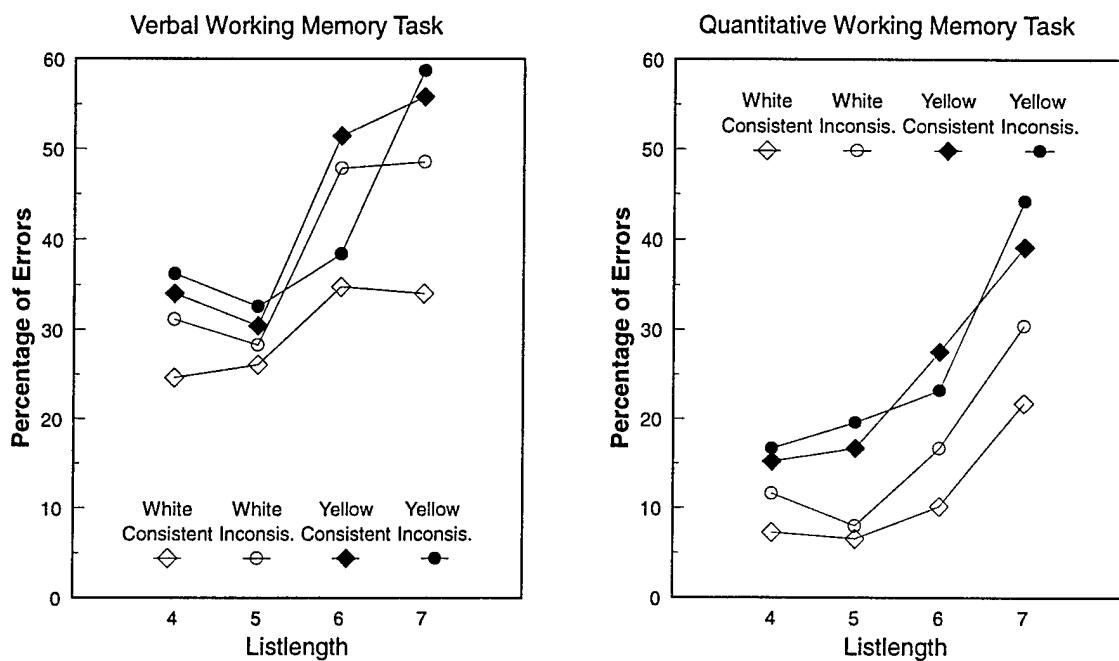


Figure 6. Mean error rate for the working memory measures by trial condition in Study 2.

Descriptive Statistics for Self-Report Measures and Speed Measures. Average scores for the anxiety measures were as follows: state anxiety ($M = 37.32$, $SD = 11.63$), trait anxiety ($M = 36.36$, $SD = 8.63$), worry ($M = 13.76$, $SD = 4.68$), and emotionality ($M = 14.53$, $SD = 5.11$). On the Cognitive Failures Questionnaire, the mean was 52.32 ($SD = 11.03$) while subjects obtained a mean latency of 1219 MS ($SD = 335$) on the verbal speed task and a mean latency of 1653 MS ($SD = 591$) on the quantitative speed task.

Spielberger, Gorsuch, Lushene, Vagg, and Jacobs (1983) reported that, among military recruits, the mean score on the state anxiety measure was 44.05 ($SD = 12.18$) and on the trait measure was 37.64 ($SD = 9.51$). Although the sample in the present study obtained a comparable score on the trait anxiety scale, they scored more than half a standard deviation below the normative sample on the state scale, which means that they may not have been anxious during the experiment. Spielberger, Gonzalez, Taylor, Anton, Algaze, Ross, and Westberry (1980)

reported that a sample of male Navy recruits obtained a mean score of 15.01 ($SD = 5.86$) on the Worry scale and a mean score of 17.05 ($SD = 5.67$) on the Emotionality scale, which places the subjects in the present study about a half a standard deviation below the normative sample on the Emotionality scale but fairly close to the normative sample on the Worry scale.

Correlations. The anxiety and working memory measures, and the CFQ and speed measures, are presented in separate tables because the correlations involving the CFQ and speed measures do not test either the cue utilization or working memory theories. The working memory measures (verbal and quantitative) were positively correlated, and all of the anxiety scales were significantly correlated (Table 2). However, only the State Anxiety scale was significantly correlated with both working memory measures.

Table 2

Correlations Between Measures of Working Memory Capacity and Anxiety

	WMQ ^a	WMV ^a	Worry	Emot ^b	Trait Anxiety	State Anxiety
(n = 69)						
WMQ ^a	-	.54**	.22	.24*	.10	.45**
WMV ^a		-	.13	.11	.12	.38**
Worry			-	.83**	.36**	.54**
Emot ^b				-	.26*	.52**
Trait Anxiety					-	.48**
State Anxiety						-

^aPercentage of Errors ^bEmotionality

*p < .05.

**p < .01.

As Table 3 shows, the CFQ (Cognitive Failures Questionnaire) was positively correlated with the Trait scale of the State-Trait Anxiety Inventory and was negatively correlated with the quantitative working memory task. Subjects who were more trait anxious tended to report more cognitive failures, but contrary to expectations, those who reported more cognitive failures made fewer errors on the quantitative working memory task.

Table 3

Correlations Involving Cognitive Failures Questionnaire and Speed Measures

	Cognitive Failures Questionnaire	Speed ^a Math	Speed ^a Verbal
WMQ ^a	-.25*	.60**	.27*
WMV ^a	-.09	.36**	.41**
Worry	.20	.29*	.29*
Emot ^b	.13	.34**	.21
Trait Anxiety	.41**	.15	.13
State Anxiety	.07	.38**	.29*
Cognitive Failures Questionnaire	-	.00	.04
Speed Math ^a	.00	-	.45**
Speed Verbal ^a	.04	.45**	-

Note. All of the correlations in this table are based on 69 subjects with the exception of the correlations with the CFQ, which are based on the 65 subjects who completed it.

^aPercentage of Errors ^bEmotionality

*p < .05. **p < .01.

^aSpeed is measured in terms of response latency.

Both of the speed measures, which were expressed in terms of response latency, (Table 3) were positively correlated with the verbal and quantitative working memory tasks, and positively correlated with the Worry scale and the State Anxiety scale. The quantitative speed measure was also positively correlated with the Emotionality scale. Worried subjects and subjects high in state anxiety were slower on the speed measures, whereas subjects with greater working memory capacity were faster.

The correlations between the individual difference measures provided only partial support for the hypothesis that anxious subjects have less working memory capacity. Only state anxiety was significantly correlated with errors on both working memory tasks. With the exception of the significant correlation between emotionality and quantitative working memory capacity, the trait anxiety, worry, and emotionality scales had nonsignificant correlations with the working memory measures.

Several individual difference variables were correlated with the percentage of undetected errors (Tables 4 and 5). Subjects who obtained higher scores on the Worry subscale of the TAI, the State Anxiety scale, and the Trait Anxiety scale made a larger number of errors. Subjects with fewer errors on the working memory measures (both verbal and quantitative) made fewer errors on the number reduction task, whereas subjects who were slower on the speed measures made a larger number of errors. However, the Emotionality subscale of the TAI and the CFQ were not significantly correlated with the percentage of errors.

Table 4

Correlations Between Percentage of Errors on Number Reduction Task, Working Memory Capacity and Anxiety

	WMQ ^a	WMV ^a	Worry	Emot ^b	Trait Anxiety	State Anxiety
(n = 69)						
All Responses	.46**	.44**	.29*	.21	.36**	.43**
Partial Match	.41**	.37**	.30*	.27*	.37**	.44**
Mismatch	.41**	.43**	.31*	.21	.39**	.43**
Old	.47**	.43**	.23	.14	.29*	.37**
WM blocks	.43**	.44**	.29*	.22	.33**	.41**
NonWM blocks	.47**	.42**	.28*	.18	.37**	.44**

^aPercentage of Errors ^bEmotionality

*p < .05. **p < .01.

Although subjects with less working memory capacity made more errors, the positive correlation between worry and errors, and between trait anxiety and errors, were probably not due to working memory limitations because the correlations between worry and working memory capacity, and between trait anxiety and working memory capacity, were not significant. Of the anxiety measures, only state anxiety was significantly correlated with both errors on the working memory task and on the number reduction task.

As Table 4 shows, subjects with greater working memory capacity made fewer errors in working memory and non working memory blocks of number reduction, and made fewer errors with all three item types (partial-match, mismatch and old). In contrast, anxious subjects, except for those who were high in emotionality, made significantly more errors under all conditions.

Table 5

Correlations Between Percentage of Errors on Number Reduction Task, Cognitive Failures Questionnaire, and Speed

	Cognitive Failures Questionnaire	Speed ^a Math	Speed ^a Verbal
All Responses	-.04	.50**	.26*
Partial Match	-.02	.46**	.30*
Mismatch	-.00	.43**	.27*
Old	-.07	.49**	.19
WM Blocks	-.06	.45**	.28*
NonWM Blocks	-.03	.52**	.23

Note. All of the correlations in this table are based on 69 subjects with the exception of the correlations with the CFQ, which are based on the 65 subjects who completed it.

*p < .05. **p < .01.

^aSpeed is measured in terms of response latency.

Multiple Regression Analyses. A principal components factor analysis of the anxiety measures revealed that all of these scales loaded on a single factor. Loadings ranged between .60

for the trait anxiety scale to .89 for the worry measure. Therefore, an anxiety factor, which was extracted from the anxiety measures, was used in the multiple regression. Predictors included the anxiety factor, a composite that was formed from the working memory measures, a composite that was formed from the speed measures, and the Cognitive Failures Questionnaire.

When the percentage of undetected errors on all of the items was examined, only the working memory composite was a significant predictor ($B = -.37$, $t = -2.91$, $p < .01$). The anxiety factor approached significance ($B = .22$, $t = 1.92$, $p = .06$) but neither the speed composite ($B = .19$, $t = 1.47$) nor the CFQ ($B = -.02$, $t = -.19$) were significant predictors. When the nonsignificant predictors were removed from the regression equation in a backward step-wise fashion, both the working memory composite ($B = -.47$, $t = -4.45$, $p < .01$) and the anxiety factor ($B = .26$, $t = 2.48$, $p < .05$) were significant and accounted for 36% of the variance in undetected errors on all items.

It appears that more than a third of the variance in undetected errors can be accounted for by working memory capacity and anxiety. This supports both the working memory and cue utilization theories that predict that anxious subjects should make more errors on the item types that were used in this study. However, the support for these theories is rather weak because the working memory composite and anxiety made independent contributions in predicting undetected errors, and as pointed out earlier, errors on the partial-match items may not have been due to the failure to attend to the last digit.

Discussion

As in Study 1, undetected errors were not significantly correlated with the CFQ, and this correlation very was small in magnitude ($< .05$). This replicates the finding of the first experiment on a larger (and somewhat different) subject population.

An important result of this study was the finding that individual differences in working memory capacity predicted undetected errors in skilled performance. Learners with greater working memory capacity made fewer undetected errors when they encountered demands for new processing sequences in near transfer conditions. This is potentially important in that technical training for real world skills rarely exposes learners to all of the conditions that they will encounter on the job. Knowing which type of learners may be less likely to make errors when they encounter new conditions in the real world could be important in jobs where errors can be disastrous.

Anxiety differences also significantly predicted transfer errors. More anxious subjects made more errors. However, neither the working memory nor the attentional explanation of how anxiety affects performance received much support in this study. First, the findings provided only weak support for the working memory theory. With the exception of state anxiety, most of the correlations between anxiety and performance on the working memory tasks were nonsignificant. Furthermore, although anxious subjects made more errors on the number reduction task, the correlations tended to be significant for all item types. Extensive practice during training with a small number of items means that practice items probably did not tax working memory capacity during transfer. Therefore, the poorer performance of anxious subjects on these items during transfer was probably not due to working memory limitations. Finally, anxiety and performance on the working memory tasks made independent contributions in predicting undetected errors, which means that the effect of anxiety on undetected errors was probably not due to working memory limitations.

The cue utilization theory predicted that anxious subjects should perform more poorly on partial-match items. Although anxious subjects made more errors with partial-match items, they also made more errors with mismatch items. The correlation between worry and errors on

partial-match items is not compelling support for the cue utilization theory because the partial-match items failed to have the desired effect. The latency of correct responses did not differ for partial-match and mismatch items, which may indicate that partial-match items taxed working memory as much as mismatch items. In addition, the latency of errors did not differ for these item types, which means that errors on partial-match items were not exclusively due to a failure to detect the difference between these items and old items.

Performance on partial-match and mismatch items may have been roughly equivalent because subjects memorized items during training (i.e., skilled performance was instance-based). Remember that only 16 instances were used in training. Informal interviews with subjects indicated that many had memorized these items. Subjects may have readily recognized that both the partial-match and mismatch items differed from the old items and slowed down for both item types.

Lack of support for the working memory theory of anxiety is probably not due to inadequate methods. The working memory tasks appeared to tax working memory capacity, and the working memory load increased the latency of correct responses on both the partial-match and mismatch new items. Furthermore, although the partial-match and mismatch items did not differ in the latency of correct responses, both item types may have taxed working memory capacity if the subjects memorized items during training.

Although the anxiety measures were significantly correlated, two interesting patterns should be noted. First, all of the anxiety measures were correlated with errors on the number reduction task except for emotionality. This is not surprising since, according to Deffenbacher (1980), many authors have found that emotionality does not reliably correlate with test performance. A more puzzling finding concerns the relationship between anxiety and working memory capacity. Only state anxiety was significantly correlated with both measures of working memory capacity. This may be due to the difference between state and trait anxiety. Both the Worry measure and the Trait Anxiety measure ask subjects how they generally feel, whereas the State Anxiety measure asks subjects how they feel right now. Perhaps how a subject feels while taking a test is more highly correlated with performance on that test than is their general feelings. As Hodges and Spielberger (1969) pointed out, anxious subjects may only perform more poorly on tasks such as digit span when state anxiety is elevated. This may also explain why the correlations between State Anxiety and errors on the number reduction task are higher than the correlations between the other measures of anxiety and performance on this task.

Study 3: Anxiety and Working Memory

Study 1 and 2 found that self-report measures of error-proneness were not related to a performance measure of error-proneness. Study 2 also found that both working memory and anxiety differences were related to skilled performance errors. However, the mechanism by which anxiety affects performance was not clear. This study was designed to further test the cue utilization and working memory theories of how anxiety affects cognitive task performance. In this experiment, all of the subjects received the same types of stimuli in training as well as the same individual difference measures. However, half of the subjects received the partial-match sequences during transfer and half received the mismatch sequences. In addition, half of each one of these groups received instructions that were designed to increase state anxiety prior to transfer.

Method

Subjects The subjects were 626 U.S. Air Force enlisted personnel who had almost completed basic training at Lackland Air Force Base, TX. There were 129 females and 497 males in the sample. Of these 626 subjects, 144 were eliminated from the study because their performance indicated a lack of effort during transfer or because data was missing for some of the transfer blocks of the skill acquisition task. The age of Air Force recruits ranges from 17 to 27.

Apparatus. The experimental task was administered on Zenith Z-248 microcomputers with standard keyboards and EGA color video monitors. Materials were presented on the monitors in 24 X 80 text mode. Software was written to achieve millisecond timing of response latency recording (Walker, 1985).

Measures. The subjects received computerized versions of three anxiety measures: the *Test Anxiety Inventory* (TAI; Spielberger, Gonzalez, Taylor, Anton, Algaze, Ross, and Westberry, 1980), the trait measure from the *State-Trait Anxiety Inventory* (STAI; Spielberger, Gorsuch, Lushene, Vagg, and Jacobs, 1983), and the *Test Anxiety Scale* (TAS; Sarason, 1980). Due to a computer error, only 370 subjects in the total sample completed the state measure from the *State-Trait Anxiety Inventory*. The subjects also completed a computerized version of the Marlowe-Crowne Social-Desirability Scale (MCSD; Crowne and Marlowe, 1964).

Only one of the working memory measures from Study 2 was used in this set of experiments. Time did not permit the administration of both measures, so the verbal working memory task was chosen because the quantitative measure was considered too easy to adequately test working memory capacity limits. In order to increase the difficulty of the quantitative measure, the rate of presentation of the items had to be increased considerably, which meant that performance may have been unduly influenced by processing speed. The verbal working memory measure in these experiments was identical to the one used in Study 2.

The skill task that was used was number reduction (Woltz et al, 1995). Since this task has been described in considerable detail, this section will mainly emphasize the differences between how it was used in Study 2 and how it was used in this study. The practice phase consisted of 30 blocks with 24 trials per block. The practice blocks contained a total of 48 unique instances, with 12 instances representing each of four rule sequences. The transfer phase contained four warm-up blocks that contained 24 practice items each, and twelve transfer blocks. Each of the twelve transfer blocks contained 16 practice trials, and 8 mismatch trials if the subject was in the experiment that tested the working memory theory, or 8 partial-match trials if they were in the experiment that tested the cue utilization theory. The transfer blocks contained the 48 original practice instances as well as 48 new instances that represented either 4 partial-match sequences or 4 mismatch sequences. All of the sequences used each of the four component rules (Contiguous, Last, Midpoint, and Same) an equal number of times in each position.

As in Study 2, the partial-match trials started with the same two rules and the same three digits as the practice trials. However, unlike Study 2, no attempt was made to ensure that the practice and partial-match trials ended in a digit that was adjacent. The mismatch items started with a different 2-rule sequence than the practice items, and mismatch items for a given sequence all started with a different set of three numbers to tax working memory. However, unlike the previous study, no attempt was made to ensure that mismatch and practice items that started with the same rule also started with a different digit because with 12 items per sequence this was not possible. Practice items for a given sequence started with a unique set of three digits so that during training subjects would not become familiar with items that started with the same three

digits and only differed on the last digit. During transfer, subjects could retake the previous item by pressing the spacebar if they thought they had made an error.

Procedure The subjects were tested in groups of 30-40, with each subject at an individual testing carrel containing a microcomputer. Each subject participated for 3.5 hours in a series of experimental tasks.

At the beginning of each session, the subjects were given a general orientation to the experiment and a few minutes of practice locating keys on the computer keyboard. All of the instructions were computer administered, and proctors were available to answer questions. Following the orientation, the subjects completed a series of tasks, which are outlined below.

At the beginning of the experiment, subjects were provided with examples of each of the four component rules that are used in the number reduction task and received four practice trials with each rule. They also saw examples of rule sequences and received three practice sequences. The subjects took 10 practice blocks of number reduction, with each block consisting of 24 items. Either the STAI and the TAS, the TAI and the social desirability scale, or the working memory task was completed next. Another 10 practice blocks of number reduction followed, with either the STAI and the TAS, the TAI and the social desirability scale, or the working memory task being presented. A third set of 10 practice blocks was again followed by either the STAI and the TAS, the TAI and the social desirability scale, or the working memory task. Subjects finished the session by taking the transfer blocks of number reduction, which consisted of 16 blocks of 24 items. Prior to the transfer blocks, half of the subjects were exposed to the anxiety manipulation. They were told that they would be tested on the number reduction task and that their performance on this test was extremely important to their career. The instructions indicated that those who don't perform well may experience difficulties in critical Air Force jobs and often do not complete their specialty training.

During the practice blocks of number reduction, feedback was provided to discourage errors and increase response speed. Latency feedback was presented for 1 sec following correct responses, and the word WRONG and a low tone was provided for 2 sec following incorrect responses. After each block of trials, the percent correct and median latency for the block was presented along with instructions for performance on the next block. If the subject's error rate was 15% or higher, they were told to slow down to reduce their errors. With an error rate of 5% or less, they were instructed to respond more quickly. Each subject was told to go faster than in the previous block and get approximately 90% of the trials correct. Following this instruction, they were shown the median latency for each previous block.

During the transfer phase of the number reduction task, subjects were told to achieve 100% accuracy while going as fast as possible. Subjects were instructed to press the spacebar to retake any trial on which they thought they had made an error. They were instructed that corrected error trials would not count against their goal of 100% accuracy. No accuracy feedback was provided during transfer, but subjects did receive latency feedback at the end of each transfer block.

Results

Number Reduction Mean Data. The learning curves (Figure 7) reveal that subjects had trouble achieving the performance goal of 10% errors during training. The mean error rate varied between 10 and 20% during the training phase. However, as in Study 2, latency declined with practice, and a power function provided a close fit to the data. Finally, when the subjects were asked to detect and correct their errors (Blocks 31 through 34), their error rate declined while latency increased.

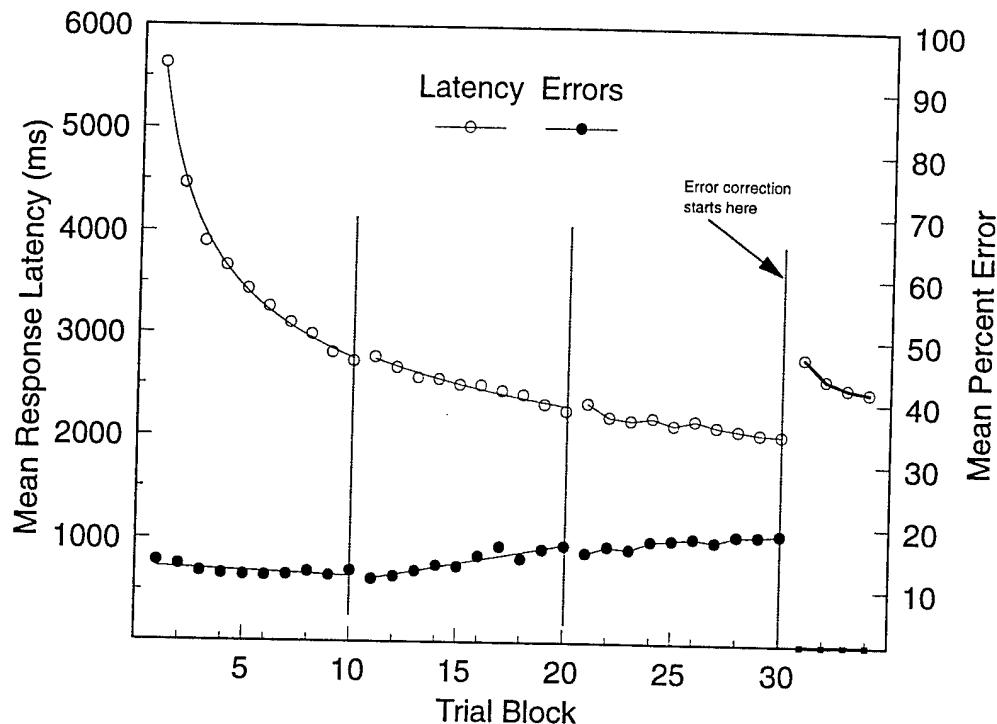


Figure 7. Mean latency and error data for training blocks in Study 3.

As in Study 2, only undetected errors are reported in this study because the vast majority of errors were undetected. The mean percentage of undetected errors was 15.77% ($SD = 16.25$) and the mean percentage of detected errors was 3.15% ($SD = 3.54$), which was significantly different ($F(1,481) = 262.18$, $MSe = 146.37$). Since none of the mean data for the number reduction task was affected by the presence of the anxiety manipulation, the effects of this manipulation will not be discussed until the regression results are presented.

In both the cue utilization and working memory experiments, subjects were slower and made more undetected errors with new items than with old items. As Figure 8 shows, in the cue utilization experiment, the percentage of undetected errors was higher on the partial-match items ($M = 19.11\%$, $SD = 18.23\%$) than on the old items ($M = 8.47\%$, $SD = 13.39\%$, $F(1,249) = 160.18$, $MSe = 89.67$). Subjects were also slower on the partial-match items ($M = 2549$ MS, $SD = 999$ MS) than on the old items ($M = 2129$ MS, $SD = 835$ MS, $F(1,249) = 142.28$, $MSe = 153436.91$).

In the working memory experiment (Figure 9), the undetected error rate on the mismatch items was higher ($M = 32.99\%$, $SD = 25.75\%$) than on the old items ($M = 13.28\%$, $SD = 17.86\%$, $F(1,229) = 160.87$, $MSe = 271.62$), and the latency of correct responses was higher for the mismatch items ($M = 3327$ MS, $SD = 1287$ MS) than for the old items ($M = 2685$ MS, $SD = 880$ MS, $F(1,229) = 167.94$, $MSe = 278596.27$).

When the latency of correct responses and the percentage of undetected errors were compared for the partial-match items in the cue utilization experiment and the mismatch items in the working memory experiment (compare Figures 8 & 9), it was found that subjects made significantly more errors on the mismatch items than on the partial-match items ($F(1,480) = 47.2$, $MSe = 490.86$), and were slower on the mismatch items than on the partial-match items ($F(1,480) = 55.43$, $MSe = 1314023$).

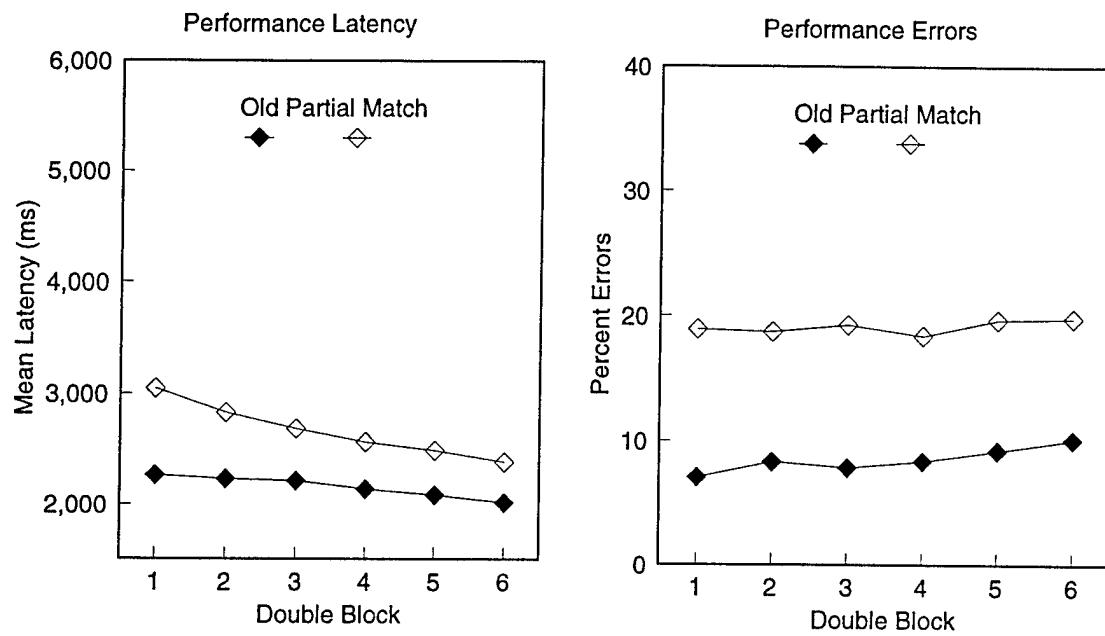


Figure 8. Mean latency and error data from transfer blocks of the cue utilization experiment.

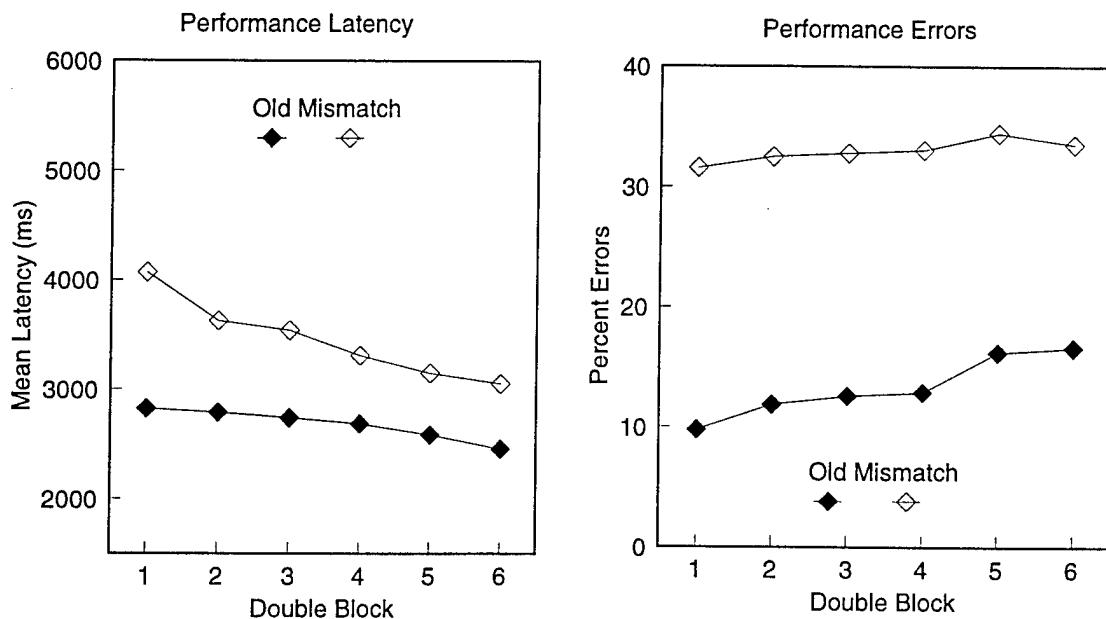


Figure 9. Mean latency and error data from transfer blocks of the working memory experiment.

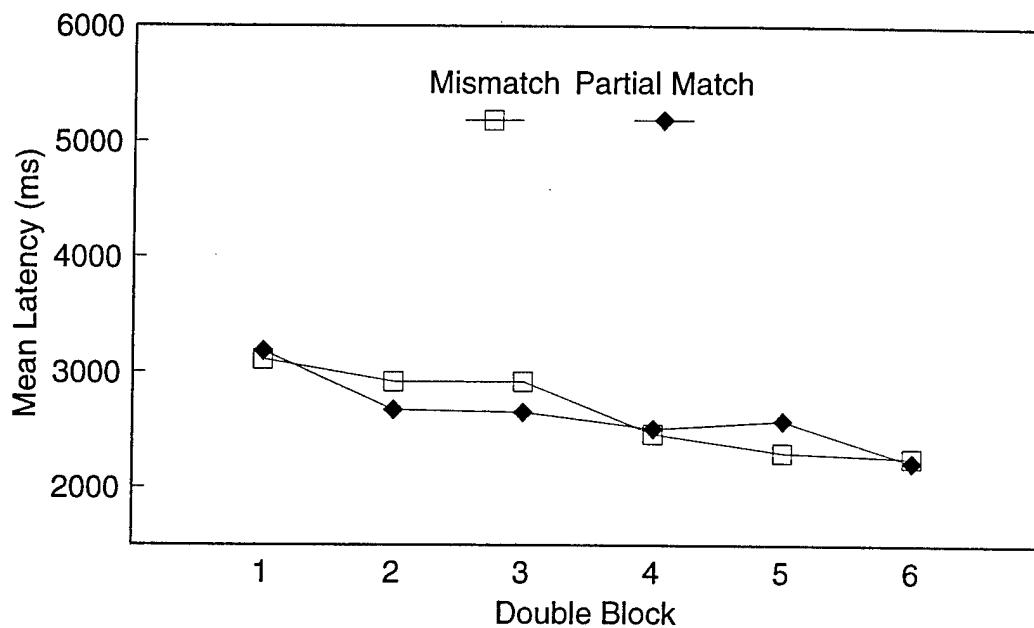


Figure 10. Mean latency of undetected errors for mismatch and partial match trials.

The results suggest that both types of new items may have taxed working memory capacity but that mismatch trials taxed it more. Subjects were slower and made more errors on the partial-match items than on the practice items, and then made more errors and were slower on the mismatch items than on the partial-match items. This is consistent with the expectation that partial-match items should tax working memory capacity less than mismatch items since partial-match items shared the same first two rules and the same first three digits with the practice items. However, as can be seen in comparing Figures 9 and 10, the latency of undetected errors on the partial-match items ($M = 2699$ MS, $SD = 1554$ MS) was not significantly different from the latency of undetected errors on the mismatch items ($M = 2653$ MS, $SD = 1518$ MS, $F(1,455) = .17$, $MSe = 2338828.9$). As in Study 2, this was inconsistent with the notion that errors on the matching items were due to a failure to detect the difference between the practice items and the matching items on the last digit.

Working Memory Task Mean Data. The effects of listlength, color of the stimuli, and consistency of color were examined separately for the working memory and cue utilization experiments. These working memory manipulations appeared to have the desired effect. As Figure 11 illustrates, accuracy declined as listlength increased for both the working memory, $F(3,228) = 90.42$, and cue utilization experiments, $F(3,248) = 69.39$. There was also a significant interaction between color and consistency for both the working memory, ($F(1,230) = 57.62$, $MSe = 894.18$) and the cue utilization, ($F(1,250) = 17.25$, $MSe = 923.75$) experiments. Overall, subjects performed better with white stimuli, which did not require any transformations. The mean percentage of errors on the working memory task in the cue utilization experiment was 46.36% ($SD = 24.13$), and in the working memory experiment was 48.65% ($SD = 24.73$).

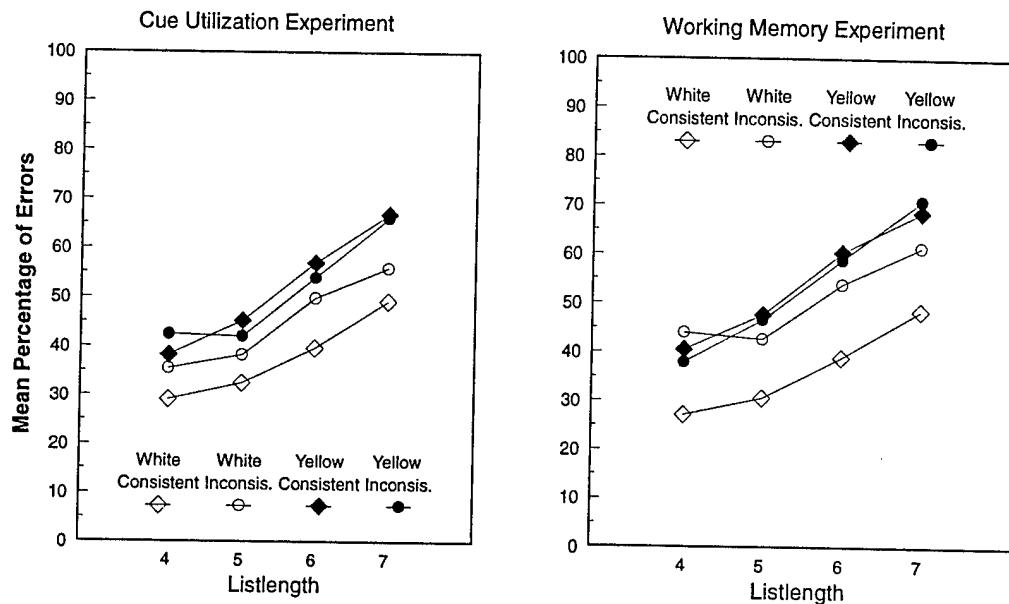


Figure 11. Mean error rate for the working memory measure in Study 3 by condition.

Descriptive Statistics for the Self-Report Measures. Average scores for the self-report measures were as follows: State Anxiety Scale ($M = 34.78$, $SD = 10.42$), Trait Anxiety Scale ($M = 34.97$, $SD = 8.64$), Test Anxiety Scale ($M = 12.24$, $SD = 7.05$), Worry scale ($M = 12.31$, $SD = 4.54$), Emotionality scale ($M = 12.79$, $SD = 4.62$) and Social Desirability Scale ($M = 21.16$, $SD = 5.51$). As reported earlier, the means for male military recruits in the normative samples are: 44.05 ($SD = 12.18$) on the State Anxiety Scale, 37.64 ($SD = 9.51$) on the Trait Anxiety Scale, 15.01 ($SD = 5.86$) on the Worry scale, and 17.05 ($SD = 5.67$) on the Emotionality scale. Male college students obtained a mean score of 15.06 ($SD = 5.58$) on the Social Desirability Scale, but no norms were available for the Test Anxiety Scale.

The subjects in this study obtained roughly the same score on the Trait Anxiety scale as the normative sample, but scored about one standard deviation lower on the State Anxiety Scale and about a half a standard deviation lower on the Worry and Emotionality scales. Although the normative sample was composed of college students, it should be noted that the recruits in my study scored one standard deviation higher on the Social Desirability Scale.

Correlations. The data will be presented separately for the cue utilization and working memory experiments. In the cue utilization study, the anxiety measures were negatively correlated with the social desirability scale (Table 6), which suggests that self-report measures of anxiety may be influenced by the need to appear in a positive light. Two of the anxiety measures were significantly correlated with the working memory measure (state anxiety and test anxiety), but state anxious subjects performed better on the working memory measure whereas test anxious subjects performed more poorly. The relatively small correlations ($< .20$) between anxiety and working memory capacity leads one to conclude that there is little support for the working memory theory of anxiety in this experiment.

No support was found for the cue utilization theory because none of the anxiety measures were significantly correlated with errors on the number reduction task (Table 7). As shown in

Table 8, this was not due to greater cautiousness. None of the anxiety measures, with the exception of worry, were correlated above .15 with the latency of correct responses on either the practice items or the matching items.

Table 6

Correlations Between Working Memory Capacity, Anxiety, and Social Desirability in Cue Utilization Experiment

	WMV ^a	Worry	Emot ^b	Trait Anxiety	State Anxiety	Test Anxiety	Social Desirability
WMV ^a	---	.12	.11	-.05	-.19*	.16**	.06
Worry		---	.82**	.53**	.33**	.80**	-.30**
Emot ^b			---	.54**	.40**	.81**	-.29**
Trait Anxiety				---	.68**	.53**	-.42**
State Anxiety					---	.41**	-.39**
Test Anxiety						---	-.27**
Social Desirability							---

^a Percentage of Errors ^b Emotionality

* p < .05

**p < .01

Note: All of the correlations in this table are based on 251 subjects except for the correlations that involve the state anxiety measure, which are based on the 148 subjects who completed it.

In the experiment that tested the working memory theory, the anxiety measures were negatively correlated with the social desirability scale (Table 9). As in the cue utilization experiment, responses on the anxiety scales may have been influenced by the need to appear in a socially desirable light. In general, the working memory experiment provided little support for the working memory theory. Although the anxiety measures were positively correlated with errors on the working memory task, the largest correlation did not exceed .30.

Furthermore, as shown in Table 10, the correlations between anxiety and the percentage of undetected errors were small (below .25). As in the cue utilization experiment, there was little evidence that anxious subjects were more cautious. None of the anxiety measures were

correlated with the latency of correct responses on either the practice items or the mismatch items (Table 11).

Table 7

Correlations Between Percentage of Errors on Number Reduction Task, Working Memory Capacity, Anxiety, and Social Desirability in Cue Utilization Experiment

	WMV ^a	Worry	Emot ^b	Trait Anxiety	State Anxiety	Test Anxiety	Social Desirability
All Responses	.23**	.00	.01	.00	.05	.09	.01
Partial Match	.19**	-.04	.00	.00	.04	.08	.06
Old	.22**	.03	.01	.00	.05	.08	-.03

^a Percentage of Errors ^b Emotionality

**p < .01

Note: All of the correlations in this table are based on 251 subjects except for the correlations that involve the state anxiety measure, which are based on the 148 subjects who completed it.

Table 8

Correlations Between Latency of Correct Responses on Number Reduction Task, Working Memory Capacity, Anxiety, and Social Desirability in Cue Utilization Experiment

	WMV ^a	Worry	Emot ^b	Trait Anxiety	State Anxiety	Test Anxiety	Social Desirability
All Responses	.17**	.24**	.13	.07	.13	.15*	.04
Partial Match	.17**	.24**	.11	.08	.10	.12	.03
Old	.15*	.23**	.13*	.06	.13	.15*	.05

Note: All of the correlations in this table are based on 251 subjects except for the correlations that involve the state anxiety measure, which are based on the 148 subjects who completed it.

^a Percentage of Errors ^b Emotionality

* p < .05

**p < .01

Table 9

Correlations Between Working Memory Capacity, Anxiety, and Social Desirability in Working Memory Experiment

	WMV ^a	Worry	Emot ^b	Trait Anxiety	State Anxiety	Test Anxiety	Social Desirability
WMV ^a	---	.30**	.28**	.21**	.21*	.24**	.02
Worry		---	.81**	.51**	.53**	.73**	-.22**
Emot ^b			---	.50**	.53**	.71**	-.30**
Trait Anxiety				---	.55**	.62**	-.42**
State Anxiety					---	.59**	-.21*
Test Anxiety						---	-.27**
Social Desirability							---

^a Percentage of Errors ^b Emotionality

* p < .05

**p < .01

Note: All of the correlations in this table are based on 231 subjects except for the correlations that involve the state anxiety measure, which are based on the 136 subjects who completed it.

Table 10

Correlations Between Percentage of Errors on Number Reduction Task, Working Memory Capacity, Anxiety, and Social Desirability in Working Memory Experiment

	WMV ^a	Worry	Emot ^b	Trait Anxiety	State Anxiety	Test Anxiety	Social Desirability
All Responses	.32**	.23**	.14*	.04	.15	.16*	.02
Mismatch	.34**	.21**	.10	.02	.16	.14*	.04
Old	.23**	.19**	.13*	.04	.10	.14*	-.01

^a Percentage of Errors ^b Emotionality

*p < .05

**p < .01

Note: All of the correlations in this table are based on 231 subjects except for the correlations that involve the state anxiety measure, which are based on the 136 subjects who completed it.

Table 11

Correlations Between Latency of Correct Responses on Number Reduction Task, Working Memory Capacity, Anxiety, and Social Desirability in Working Memory Experiment

	WMV ^a	Worry	Emot ^b	Trait Anxiety	State Anxiety	Test Anxiety	Social Desirability
All Responses	-.03	.02	.01	.04	.05	.05	.04
Mismatch	-.07	.03	-.01	.00	.02	.02	.06
Old	.01	.01	.02	.06	.06	.06	.02

^a Percentage of Errors ^b Emotionality

* p < .05

**p < .01

Note: All of the correlations in this table are based on 231 subjects except for the correlations that involve the state anxiety measure, which are based on the 136 subjects who completed it.

The correlations between errors on the number reduction task and errors on the working memory task reveal that both types of new items may have taxed working memory capacity. Significant correlations between the percentage of errors on the working memory task and errors on matching items (Table 7) and mismatch items (Table 10) indicate that both types of new items taxed working memory capacity. This supports the conclusions that were drawn from examining the number reduction mean data.

Multiple Regression Analyses. The results of a set of regression runs will be presented separately for the working memory and cue utilization experiments. A principal components factor analysis of the anxiety measures revealed that all of these scales loaded on a single factor. Loadings in the working memory experiment ranged between .74 for the trait anxiety measure and .90 for the worry measure, while in the cue utilization study the loadings ranged between .72 for the trait anxiety measure and .92 for the Emotionality scale. Therefore, an anxiety factor, which was extracted from the anxiety measures, was used in the regression analyses.

Predictors in the regressions included the anxiety factor, the working memory measure, whether the subjects had been exposed to the anxiety manipulation, and the interaction between the anxiety factor and the anxiety manipulation. The interaction between the anxiety factor and the anxiety manipulation was included to examine whether subjects who scored higher on the anxiety measures made more errors when exposed to the anxiety manipulation. The social desirability scale was not included because this scale was used to simply check on the validity of the self-report measures of anxiety, and no predictions had been made relating social desirability to errors on the skill task.

In the working memory experiment, both the anxiety manipulation and the working memory task were significant predictors of undetected errors on the practice items (*B* for the

working memory task = -.22, $t = -3.23$, $p < .01$; B for the anxiety manipulation = $-.18$, $t = -2.76$, $p = .01$) but the anxiety factor was nonsignificant ($B = .09$, $t = .95$) and the interaction between the anxiety factor and the anxiety manipulation was not significant ($B = -.01$, $t = -.06$). For the mismatch items, only the working memory task was significant ($B = -.33$, $t = -5.10$, $p < .01$). Subjects who performed more poorly on the working memory measure made more undetected errors on all item types, which indicates that performance on the skill task is related to working memory capacity. However, no support was found for the working memory theory which predicted that anxiety would increase the number of errors on mismatch items. The anxiety manipulation actually decreased the number of undetected errors on practice items which was opposite to its intended effect, and did not interact with the anxiety factor in producing errors. When the nonsignificant predictors were removed, 8% of the variance in undetected errors on practice items was accounted for by working memory capacity ($B = -.24$, $t = -3.82$, $p < .01$) and the anxiety manipulation ($B = -.17$, $t = -2.74$, $p = .01$), and 11% of the variance in undetected errors on mismatch items could be accounted for by working memory capacity ($B = -.34$, $t = -5.43$, $p < .01$).

The cue utilization experiment provided a similar pattern of results. On the practice items, both working memory capacity and the anxiety manipulation were significant predictors of undetected errors (B for working memory capacity = $-.22$, $t = -3.56$, $p < .01$; B for the anxiety manipulation = $-.15$, $t = -2.40$, $p < .05$) but the anxiety factor ($B = .00$, $t = -.03$) and the interaction between the anxiety factor and the anxiety manipulation ($B = .05$, $t = .44$) were not significant predictors. On the matching items both the anxiety manipulation and the working memory measure were significant predictors (B for the working memory measure = $-.19$, $t = -3.11$, $p < .01$; B for the anxiety manipulation = $-.18$, $t = -2.82$, $p < .01$) but the anxiety factor ($B = -.06$, $t = -.55$) and the interaction between the anxiety factor and the anxiety manipulation ($B = .10$, $t = .88$) were not. As in the working memory experiment, anxiety was not a predictor of undetected errors, which means that the cue utilization theory was not supported by the data in this experiment. Again, working memory capacity was a significant predictor of undetected errors on all item types, the anxiety manipulation was related to undetected errors but in the wrong direction, and the interaction between the anxiety factor and the anxiety manipulation was not significant. When nonsignificant predictors were removed from the regression equation, 7% of the variance in undetected errors on practice items could be accounted for by working memory capacity ($B = -.22$, $t = -3.65$, $p < .01$) and the anxiety manipulation ($B = -.14$, $t = -2.37$, $p < .05$). Working memory capacity ($B = -.19$, $t = -3.15$, $p < .01$) and the anxiety manipulation ($B = -.18$, $t = -2.90$, $p < .01$) accounted for 7% of the variance in undetected errors on matching items.

Discussion

This experiment replicated the Study 2 finding that differences in working memory capacity are related to the propensity to make undetected errors in near transfer conditions of a cognitive skill. However, there was little support for either the cue utilization or working memory theory of how anxiety affects performance. Although the cue utilization theory may not have been adequately tested because the matching items did not have the desired effect on performance, inadequate methods do not explain the lack of support for the working memory theory of anxiety. The mismatch items tapped working memory capacity, and a working memory measure was included which appeared to tax working memory capacity.

The results even question the more basic notion that anxiety affects performance. Anxiety was not very highly correlated with working memory capacity and was not a significant predictor of undetected errors. It is possible that more severe anxiety manipulations are needed to

adequately test the hypothesized anxiety-performance relationships. This study attempted to increase performance-related anxiety in half of the sample, but this manipulation failed to have the desired effect.

Conclusions and Implications

Three main conclusions can be drawn from the evidence presented in this report. First, our evidence suggests that self-report measures of error-proneness (the *Cognitive Failures Questionnaire*, and the *Error-proneness Questionnaire*) are not related to performance errors in either working memory tests or skilled performance transfer conditions. In contrast, they were more related to other self-report measures, including a measure of social desirability. This leads us to believe that these self-report measures tell us more about how a person perceives themselves than what a person's true error-proneness is in performance situations.

Second, our evidence suggests that individual differences in working memory capacity is a significant predictor of skilled performance errors. This has implications for both applied and theoretical researchers. In applied settings, being able to predict which individuals would be most susceptible to undetected performance slips even after extensive training may someday have important practical implications in critical military jobs. We recommend that further research be conducted on this issue, especially with complex and realistic skill tasks.

With respect to theoretical research on individual differences in skill acquisition, the degree to which general ability differences can predict later performance has been a topic of some controversy. Some have argued it is difficult to predict individual differences in the late phases of skill acquisition with general ability measures (e.g., Fleishman, 1972). Findings supporting this position generally involve task environments with consistent mappings between stimuli and responses. Others (Ackerman, 1992) have demonstrated that general ability can continue to predict post training performance even after extensive practice, so long as the task environment involves inconsistent mappings.

Our research is unique with respect to this issue in that it has investigated performance in very near transfer performance. That is, we studied performance differences under task demands that differed only slightly from training conditions - so slightly that subjects were typically unaware of the differences. Thus, our transfer conditions involved some inconsistencies with respect to prior skill practice, but these inconsistencies were very subtle. We believe that this form of transfer is highly relevant to many real world-training environments. Technical training courses may train personnel on a skill under a variety of task conditions. However, when individuals begin to perform the skill in the real world, they typically encounter a greater variety of conditions that may differ slightly from those practiced during training. In such near transfer environments we found that working memory, a general intellectual ability, still predicted individual differences in error-making after extensive training. This is consistent with Ackerman's theory (1987, 1988, 1992) about the prediction of post practice performance - namely that the ability to predict post practice performance depends on both the consistency of the task relative to training and the kinds of measures used to predict performance. However, our findings extend this theory's applicability to very near transfer conditions and the undetected performance errors that can occur therein.

The third major conclusion from this research concerns individual differences in anxiety. Here our evidence was mixed, so we can only conclude that more research is needed on this topic. We found some evidence that more anxious individuals make more undetected errors in near transfer skilled performance. This was consistent with several theoretical explanations of how anxiety may limit the efficiency of cognitive performance. However, we failed to replicate

this finding under slightly different experimental conditions. We recommend that this issue be studied further, especially under anxiety provoking performance conditions.

References

Ackerman, P. L. (1987). Individual differences in skill learning: An integration of psychometric and information processing perspectives. *Psychological Bulletin, 102*, 3-27.

Ackerman, P. L. (1988). Determinants of individual differences during skill acquisition: Cognitive abilities and information processing. *Journal of Experimental Psychology: General, 117*, 288-318.

Ackerman, P. L. (1992). Predicting individual differences in complex skill acquisition: Dynamics of ability determinants. *Journal of Applied Psychology, 5*, 598-614.

Anderson, J.R. (1983). *The architecture of cognition*. Cambridge, MA:Harvard University Press.

Anderson, J.R. (1987). Skill acquisition: Compilation of weak-method problem solutions. *Psychological Review, 94*, 192-210.

Bell, B.G., Gardner, M.K., & Woltz, D.J. (in press). Individual differences in skilled performance. *Learning and Individual Differences*.

Broadbent, D.E., Cooper, P.F., FitzGerald, P., & Parkes, K.R. (1982). The Cognitive Failures Questionnaire (CFQ) and Its Correlates. *British Journal of Clinical Psychology, 21*, 1-16.

Bryan, W.L. & Harter, N. (1899). Studies on the telegraphic language: The acquisition of a hierarchy of habits. *Psychological Review, 6*, 345-375.

Crossman, E. R. F. W. (1959). A theory of the acquisition of speed-skill. *Ergonomics, 2*, 153-166.

Crowne, D. P., & Marlow, D. (1964). *The approval motive: Studies in evaluative dependence*. New York: Wiley.

Deffenbacher, J. L., (1980). Worry and emotionality in test anxiety. In I. G. Sarason (Ed.), *Test anxiety: Theory, research, and applications* (pp.111-128). Hillsdale, NJ: Erlbaum.

Dickman, S. J. (1990). Functional and Dysfunctional Impulsivity: Personality and Cognitive Correlates. *Journal of Personality and Social Psychology, 58*, 95-102.

Eysenck, M.W. (1983). Anxiety and Individual Differences. In G.R.J. Hockey (Ed.), *Stress and Fatigue in Human Performance* (pp. 273-298). Chichester: John Wiley and Sons Ltd.

Eysenck, M.W. (1984). *A Handbook of Cognitive Psychology*. London: Lawrence Erlbaum Associates.

Hodges, W.F. & Spielberger, C.D. (1969). Digit Span: An Indicant of Trait or State Anxiety? *Journal of Consulting and Clinical Psychology, 33*, 430-434.

Jackson, D. N. (1976). *Jackson Personality Inventory Test Manual*. London, Ontario. Sigma Assessment Systems, Inc.

Jackson, D. N. (1989). *Personality Research Form Manual*. London, Ontario. Sigma Assessment Systems, Inc.

Kyllonen, P. C., Woltz, D. J., Christal, R. E., Tirre, W. C., Shute, V. J., & Chaiken, S. (1990). *CAM-4: Computerized battery of cognitive ability tests*. Unpublished computer program, Brooks AFB, TX: Armstrong Laboratory.

Reason, J. T. (1990). *Human error*. Cambridge, UK: Cambridge University Press.

Reason, J.T., & Mycielska, K. *Absent-Minded? The Psychology of Mental Lapses and Everyday Errors*. Englewood Cliffs, N.J.: Prentice-Hall, 1982.

Sarason, I.G. (1980). Introduction to the Study of Test Anxiety. In I. G. Sarason (Ed.), *Test Anxiety: theory, research, and applications*. (pp. 3-14). Hillsdale: Lawrence Erlbaum and Associates.

Spielberger, C.D. (1972). Conceptual and Methodological Issues in Anxiety Research. In C.D. Spielberger (Ed.), *Anxiety: Current trends in theory and research*. Vol. 2. (pp. 481-493). New York: Academic Press.

Spielberger, C.D., Gonzalez, H.P., Taylor, C.J., Anton, W.D., Algaze, B., Ross, G.R., & Westberry, L.G. (1980). *Manual for the Test Anxiety Inventory*. Palo Alto, Ca.: Consulting Psychologists Press.

Spielberger, C.D., Gorsuch, R.L., Lushene, R., Vagg, P.R., & Jacobs, G.A. (1983). *Manual for the State-Trait Anxiety Inventory*. Palo Alto, CA.: Consulting Psychologists Press.

Thurstone, L. L., & Thurstone, T. G. (1941). *Factorial studies of intelligence*. Chicago: University of Chicago Press.

Walker, R. (1985). *PLATS: Software for Cognitive Tasks*. Unpublished computer program.

Woltz, D. J., Gardner, M. K., & Bell, B.G., (1995). Undetected Errors in Sequential Cognitive Skills. *Journal of Experimental Psychology: General* (Manuscript Submitted for Publication).

Woltz, D. J., Bell, B.G., Kyllonen, P.C., & Gardner, M.K. Memory for Order of Operations in the Acquisition and Transfer of Sequential Cognitive Skills. *Journal of Experimental Psychology: Learning, Memory, and Cognition* (1996).

Graduate Students Supported

Brian G. Bell

Ph.D. dissertation title: *The effects of anxiety on skilled performance errors*

Expected Graduation Date: June, 1997.

List of Publications

Bell, B. G., Gardner, M. K., & Woltz, D. J. (in press). Individual differences in undetected errors in skilled cognitive performance. *Learning and Individual Differences*.

Larkin, A. A. , Woltz, D. J., Reynolds, R. E., & Clark, E. (1996). Conceptual priming differences and reading ability. *Contemporary Educational Psychology, 21*, 279-303.

Woltz, D. J. (1996). Perceptual and conceptual priming in a semantic reprocessing task. *Memory & Cognition, 24*, 429-440.

Woltz, D. J., Bell, B. G., Kyllonen, P. C., & Gardner, M. K. (1996). Memory for order of operations in the acquisition and transfer of sequential cognitive skills. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 22*, 438-457.

Woltz, D. J., & Shute, V. J. (1995). Time course of forgetting exhibited in repetition priming of semantic comparisons. *American Journal of Psychology, 108*, 499-525.

Ackerman, P. A., & Woltz, D. J. (1994). Determinants of learning and performance in an associative memory/substitution task: Task constraints, individual differences, volition, and motivation. *Journal of Educational Psychology, 86*, 1-29.

Gawlick-Grendall, L. A., & Woltz, D. J. (1994). Dominance of meanings for 120 homographs. *Behavior Research Methods, Instruments, & Computers, 26*, 5-25.

Woltz, D. J., & Shute, V. J. (1993). Individual difference in repetition priming and its relationship to declarative knowledge acquisition. *Intelligence, 17*, 333-359.

List of Presentations

Bell, B. G., Gardner, M. K., & Woltz, D. J. (1996, April). *Anxiety and skilled performance errors*. Paper presented at the annual meeting of the Rocky Mountain Psychological Association, Park City, Utah.

Schliebner, S., Woltz, D. J., Gardner, M. K., Bell, B. G., & Scharine, A. (1996, April). *Processing sequence memory and the modifiability of highly skilled memory*. Paper presented at the annual meeting of the Rocky Mountain Psychological Association, Park City, Utah.

Gardner, M. K., Bell, B. G., & Woltz, D. J. (1995, June). *Individual differences in undetected errors in skilled performance*. Paper presented at the annual meeting of the American Psychological Society, New York.

Madsen, J. G., & Woltz, D. J. (1995, April). *Changes in access to verbal knowledge as procedural skill develops*. Paper presented at the annual meeting of the American Educational Research Association, San Francisco.

Woltz, D. J., Larkin, A. A., Reynolds, R. E., & Clark, E. (1995, April). *Individual differences in semantic and repetition priming and their relationship to children's reading ability*. Paper presented at the annual meeting of the American Educational Research Association, San Francisco.

Numaguchi, G., Kircher, J. C., Craig, R. A., Raskin, D. C., Woltz, D. J., & Packard, R. E. (1994, October). *The Erlanger method for measuring cardiovascular activity: Correlations with blood volume and arterial pressure*. Paper presented at the annual meeting of the Society for Psychophysiological Research, Atlanta, Georgia.

Thomas, L. A., & Woltz, D. J. (1994, August). *A construct validity investigation of time orientation*. Paper presented at the annual meeting of the American Psychological Association, Los Angeles, California.

Woltz, D. J., Gardner, M. K., Bell, B. G., & Farnham, J. (1994, July). *Memory for processing sequences can produce undetected errors in skilled performance*. Paper presented at the annual meeting of the American Psychological Society, Washington, D.C..

Woltz, D. J., Gardner, M. K., & Bell, B. G. (1994, April). *Knowledge compilation and undetected errors by skilled performers*. Paper presented at the annual meeting of the American Educational Research Association, New Orleans, Louisiana.

Gardner, M. K., Woltz, D. J., & Bell, B. G. (1993, August). *The cost of expertise: Undetected errors in cognitive skill acquisition*. Invited address to U.S. Air Force Armstrong Laboratory, Brooks AFB, TX.

Woltz, D. J. (1993, June). *Individual differences in implicit and explicit memory measures across content domains*. Paper presented at the annual meeting of the American Psychological Society, Chicago, Illinois.

Bell, B. G., Woltz, D. J., & Madsen, J. G. (1993, April). *The role of memory for procedural sequences in skill acquisition*. Paper presented at the annual meeting of the Western Psychological Association, Phoenix, Arizona.